

PCT09

RAW SEQUENCE LISTING

DATE: 10/29/2001

PATENT APPLICATION: US/09/856,927

TIME: 13:21:11

Input Set : A:\Nih382-1.app

Output Set: N:\CRF3\10292001\I856927.raw

3 <110> APPLICANT: Dean, Michael
4 Allikmets, Rando
5 Bates, Susan E.
6 Fojo, Antonio T.
7 The Government of the United States of America
8 as represented by the Secretary of the
9 Department of Health and Human Services
11 <120> TITLE OF INVENTION: A Novel ATP-Binding Cassette Protein Responsible for
12 Cytotoxin Resistance
14 <130> FILE REFERENCE: 015280-382100US
16 <140> CURRENT APPLICATION NUMBER: US 09/856,927
17 <141> CURRENT FILING DATE: 2001-05-29
19 <150> PRIOR APPLICATION NUMBER: US 60/110,473
20 <151> PRIOR FILING DATE: 1998-11-30
22 <150> PRIOR APPLICATION NUMBER: WO PCT/US99/28107
23 <151> PRIOR FILING DATE: 1999-11-24
25 <160> NUMBER OF SEQ ID NOS: 6
27 <170> SOFTWARE: PatentIn Ver. 2.1
29 <210> SEQ ID NO: 1
30 <211> LENGTH: 2719
31 <212> TYPE: DNA
32 <213> ORGANISM: Homo sapiens
34 <220> FEATURE:
35 <221> NAME/KEY: CDS
36 <222> LOCATION: (205)..(2172)
37 <223> OTHER INFORMATION: MXR1
39 <400> SEQUENCE: 1
40 tttaggaacg caccgtgcac atgcttggtg gtcttggttaa gtggaaactg ctgctttaga 60
42 gtttggttgg aaggtccggg tgactcatcc caacatttac atccttaatt gttaaagcgc 120
44 tgccctccgag cgcacgcac ctgagatcct gagcctttgg ttaagaccga gctctattaa 180
46 gctgaaaaga taaaaactct ccag atg tct tcc agt aat gtc gaa gtt ttt 231
47 Met Ser Ser Ser Asn Val Glu Val Phe
48 1 5
50 atc cca gtg tca caa gga aac acc aat ggc ttc ccc gcg aca gtt tcc 279
51 Ile Pro Val Ser Gln Gly Asn Thr Asn Gly Phe Pro Ala Thr Val Ser
52 10 15 20 25
54 aat gac ctg aag gca ttt act gaa gga gct gtg tta agt ttt cat aac 327
55 Asn Asp Leu Lys Ala Phe Thr Glu Gly Ala Val Leu Ser Phe His Asn
56 30 35 40
58 atc tgc tat cga gta aaa ctg aag agt ggc ttt cta cct tgt cga aaa 375
59 Ile Cys Tyr Arg Val Lys Leu Lys Ser Gly Phe Leu Pro Cys Arg Lys
60 45 50 55
62 cca gtt gag aaa gaa ata tta tcg aat atc aat ggg atc atg aaa cct 423
63 Pro Val Glu Lys Glu Ile Leu Ser Asn Ile Asn Gly Ile Met Lys Pro
64 60 65 70
66 ggt ctc aac gcc atc ctg gga ccc aca ggt gga ggc aaa tct tcg tta 471
67 Gly Leu Asn Ala Ile Leu Gly Pro Thr Gly Gly Gly Lys Ser Ser Leu

ENTERED

RAW SEQUENCE LISTING

DATE: 10/29/2001

PATENT APPLICATION: US/09/856,927

TIME: 13:21:11

Input Set : A:\Nih382-1.app

Output Set: N:\CRF3\10292001\I856927.raw

68	75	80	85	
70	tta gat gtc tta gct gca agg aaa gat cca agt gga tta tct gga gat	519		
71	Leu Asp Val Leu Ala Ala Arg Lys Asp Pro Ser Gly Leu Ser Gly Asp			
72	90 95 100 105			
74	gtt ctg ata aat gga gca ccg cga cct gcc aat ttc aaa tgt aat tca	567		
75	Val Leu Ile Asn Gly Ala Pro Arg Pro Ala Asn Phe Lys Cys Asn Ser			
76	110 115 120			
78	ggt tac gtg gta caa gat gat gtt gtg atg ggc act ctg acg gtg aga	615		
79	Gly Tyr Val Val Gln Asp Asp Val Val Met Gly Thr Leu Thr Val Arg			
80	125 130 135			
82	gaa aac tta cag ttc tca gca gct ctt cgg ctt gca aca act atg acg	663		
83	Glu Asn Leu Gln Phe Ser Ala Ala Leu Arg Leu Ala Thr Thr Met Thr			
84	140 145 150			
86	aat cat gaa aaa aac gaa cgg att aac agg gtc att gaa gag tta ggt	711		
87	Asn His Glu Lys Asn Glu Arg Ile Asn Arg Val Ile Glu Glu Leu Gly			
88	155 160 165			
90	ctg gat aaa gtg gca gac tcc aag gtt gga act cag ttt atc cgt ggt	759		
91	Leu Asp Lys Val Ala Asp Ser Lys Val Gly Thr Gln Phe Ile Arg Gly			
92	170 175 180 185			
94	gtg tct gga gga gaa aga aaa agg act agt ata gga atg gag ctt atc	807		
95	Val Ser Gly Gly Glu Arg Lys Arg Thr Ser Ile Gly Met Glu Leu Ile			
96	190 195 200			
98	act gat cct tcc atc ttg tcc ttg gat gag cct aca act ggc tta gac	855		
99	Thr Asp Pro Ser Ile Leu Ser Leu Asp Glu Pro Thr Thr Gly Leu Asp			
100	205 210 215			
102	tca agc aca gca aat gct gtc ctt ttg ctc ctg aaa agg atg tct aag	903		
103	Ser Ser Thr Ala Asn Ala Val Leu Leu Leu Leu Lys Arg Met Ser Lys			
104	220 225 230			
106	cag gga cga aca atc atc ttc tcc att cat cag cct cga tat tcc atc	951		
107	Gln Gly Arg Thr Ile Ile Phe Ser Ile His Gln Pro Arg Tyr Ser Ile			
108	235 240 245			
110	ttc aag ttg ttt gat agc ctc acc tta ttg gcc tca gga aga ctt atg	999		
111	Phe Lys Leu Phe Asp Ser Leu Thr Leu Leu Ala Ser Gly Arg Leu Met			
112	250 255 260 265			
114	ttc cac ggg cct gct cag gag gcc ttg gga tac ttt gaa tca gct ggt	1047		
115	Phe His Gly Pro Ala Gln Glu Ala Leu Gly Tyr Phe Glu Ser Ala Gly			
116	270 275 280			
118	tat cac tgt gag gcc tat aat aac cct gca gac ttc ttc ttg gac atc	1095		
119	Tyr His Cys Glu Ala Tyr Asn Asn Pro Ala Asp Phe Phe Leu Asp Ile			
120	285 290 295			
122	att aat gga gat tcc act gct gtg gca tta aac aga gaa gaa gac ttt	1143		
123	Ile Asn Gly Asp Ser Thr Ala Val Ala Leu Asn Arg Glu Glu Asp Phe			
124	300 305 310			
126	aaa gcc aca gag atc ata gag cct tcc aag cag gat aag cca ctc ata	1191		
127	Lys Ala Thr Glu Ile Ile Glu Pro Ser Lys Gln Asp Lys Pro Leu Ile			
128	315 320 325			
130	gaa aaa tta gcg gag att tat gtc aac tcc tcc ttc tac aaa gag aca	1239		
131	Glu Lys Leu Ala Glu Ile Tyr Val Asn Ser Ser Phe Tyr Lys Glu Thr			
132	330 335 340 345			

RAW SEQUENCE LISTING

DATE: 10/29/2001

PATENT APPLICATION: US/09/856,927

TIME: 13:21:11

Input Set : A:\Nih382-1.app

Output Set: N:\CRF3\10292001\I856927.raw

```

134 aaa gct gaa tta cat caa ctt tcc ggg ggt gag aag aag aag aag atc 1287
135 Lys Ala Glu Leu His Gln Leu Ser Gly Gly Glu Lys Lys Lys Lys Ile
136          350          355          360
138 aca gtc ttc aag gag atc agc tac acc acc tcc ttc tgt cat caa ctc 1335
139 Thr Val Phe Lys Glu Ile Ser Tyr Thr Thr Ser Phe Cys His Gln Leu
140          365          370          375
142 aga tgg gtt tcc aag cgt tca ttc aaa aac ttg ctg ggt aat ccc cag 1383
143 Arg Trp Val Ser Lys Arg Ser Phe Lys Asn Leu Leu Gly Asn Pro Gln
144          380          385          390
146 gcc tct ata gct cag atc att gtc aca gtc gta ctg gga ctg gtt ata 1431
147 Ala Ser Ile Ala Gln Ile Ile Val Thr Val Val Leu Gly Leu Val Ile
148          395          400          405
150 ggt gcc att tac ttt ggg cta aaa aat gat tct act gga atc cag aac 1479
151 Gly Ala Ile Tyr Phe Gly Leu Lys Asn Asp Ser Thr Gly Ile Gln Asn
152 410          415          420          425
154 aga gct ggg gtt ctc ttc ctg acg acc aac cag tgt ttc agc agt 1527
155 Arg Ala Gly Val Leu Phe Phe Leu Thr Asn Gln Cys Phe Ser Ser
156          430          435          440
158 gtt tca gcc gtg gaa ctc ttt gtg gta gag aag aag ctc ttc ata cat 1575
159 Val Ser Ala Val Glu Leu Phe Val Val Glu Lys Lys Leu Phe Ile His
160          445          450          455
162 gaa tac atc agc gga tac tac aga gtg tca tct tat ttc ctt gga aaa 1623
163 Glu Tyr Ile Ser Gly Tyr Tyr Arg Val Ser Ser Tyr Phe Leu Gly Lys
164          460          465          470
166 ctg tta tct gat tta tta ccc atg agg atg tta cca agt att ata ttt 1671
167 Leu Leu Ser Asp Leu Leu Pro Met Arg Met Leu Pro Ser Ile Ile Phe
168          475          480          485
170 acc tgt ata gtg tac ttc atg tta gga ttg aag cca aag gca gat gcc 1719
171 Thr Cys Ile Val Tyr Phe Met Leu Gly Leu Lys Pro Lys Ala Asp Ala
172 490          495          500          505
174 ttc ttc gtt atg atg ttt acc ctt atg atg gtg gct tat tca gcc agt 1767
175 Phe Phe Val Met Met Phe Thr Leu Met Met Val Ala Tyr Ser Ala Ser
176          510          515          520
178 tcc atg gca ctg gcc ata gca gca ggt cag agt gtg gtt tct gta gca 1815
179 Ser Met Ala Leu Ala Ile Ala Ala Gly Gln Ser Val Val Ser Val Ala
180          525          530          535
182 aca ctt ctc atg acc atc tgt ttt gtg ttt atg atg att ttt tca ggt 1863
183 Thr Leu Leu Met Thr Ile Cys Phe Val Phe Met Met Ile Phe Ser Gly
184          540          545          550
186 ctg ttg gtc aat ctc aca acc att gca tct tgg ctg tca tgg ctt cag 1911
187 Leu Leu Val Asn Leu Thr Thr Ile Ala Ser Trp Leu Ser Trp Leu Gln
188          555          560          565
190 tac ttc agc att cca cga tat gga ttt acg gct ttg cag cat aat gaa 1959
191 Tyr Phe Ser Ile Pro Arg Tyr Gly Phe Thr Ala Leu Gln His Asn Glu
192 570          575          580          585
194 ttt ttg gga caa aac ttc tgc cca gga ctc aat gca aca gga aac aat 2007
195 Phe Leu Gly Gln Asn Phe Cys Pro Gly Leu Asn Ala Thr Gly Asn Asn
196          590          595          600
198 cct tgt aac tat gca aca tgt act ggc gaa gaa tat ttg gta aag cag 2055

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/856,927

DATE: 10/29/2001

TIME: 13:21:11

Input Set : A:\Nih382-1.app

Output Set: N:\CRF3\10292001\I856927.raw

```

199 Pro Cys Asn Tyr Ala Thr Cys Thr Gly Glu Glu Tyr Leu Val Lys Gln
200          605          610          615
202 ggc atc gat ctc tca ccc tgg ggc ttg tgg aag aat cac gtg gcc ttg 2103
203 Gly Ile Asp Leu Ser Pro Trp Gly Leu Trp Lys Asn His Val Ala Leu
204          620          625          630
206 gct tgt atg att gtt att ttc ctc aca att gcc tac ctg aaa ttg tta 2151
207 Ala Cys Met Ile Val Ile Phe Leu Thr Ile Ala Tyr Leu Lys Leu Leu
208          635          640          645
210 ttt ctt aaa aaa tat tct taaatttccc cttaattcag tatgatttat 2199
211 Phe Leu Lys Lys Tyr Ser
212 650          655
214 cctcacataa aaaagaagca ctttgattga agtattcaat caagtttttt tgttgttttc 2259
216 tgttcccttg ccatcacact gttgcacagc agcaattggt ttaaagagat acatttttag 2319
218 aaatcacaaac aaactgaatt aaacatgaaa gaaccaaga catcatgtat cgcatttag 2379
220 ttaatctcct cagacagtaa ccatggggaa gaaatctggt ctaatttatt aatctaaaaa 2439
222 aggagaattg aattctggaa actcctgaca agttattact gtctctggca tttgtttcct 2499
224 catctttaaa atgaataggt aggttagtag cccttcagtc ttaatacttt atgatgctat 2559
226 ggtttgccat tattttaatat atgacaaatg tattaatgct atactggaaa tgtaaaattg 2619
228 aaaatatggt gaaaaaaaga ttctgtctta tagggtaaaa aaagccaccg gtgatagaaa 2679
230 aaaaatcttt ttgataagca cattaaagtt aatagaactt 2719
233 <210> SEQ ID NO: 2
234 <211> LENGTH: 655
235 <212> TYPE: PRT
236 <213> ORGANISM: Homo sapiens
238 <400> SEQUENCE: 2
239 Met Ser Ser Ser Asn Val Glu Val Phe Ile Pro Val Ser Gln Gly Asn
240 1          5          10          15
242 Thr Asn Gly Phe Pro Ala Thr Val Ser Asn Asp Leu Lys Ala Phe Thr
243          20          25          30
245 Glu Gly Ala Val Leu Ser Phe His Asn Ile Cys Tyr Arg Val Lys Leu
246          35          40          45
248 Lys Ser Gly Phe Leu Pro Cys Arg Lys Pro Val Glu Lys Glu Ile Leu
249          50          55          60
251 Ser Asn Ile Asn Gly Ile Met Lys Pro Gly Leu Asn Ala Ile Leu Gly
252 65          70          75          80
254 Pro Thr Gly Gly Gly Lys Ser Ser Leu Leu Asp Val Leu Ala Ala Arg
255          85          90          95
257 Lys Asp Pro Ser Gly Leu Ser Gly Asp Val Leu Ile Asn Gly Ala Pro
258          100          105          110
260 Arg Pro Ala Asn Phe Lys Cys Asn Ser Gly Tyr Val Val Gln Asp Asp
261          115          120          125
263 Val Val Met Gly Thr Leu Thr Val Arg Glu Asn Leu Gln Phe Ser Ala
264          130          135          140
266 Ala Leu Arg Leu Ala Thr Thr Met Thr Asn His Glu Lys Asn Glu Arg
267 145          150          155          160
269 Ile Asn Arg Val Ile Glu Glu Leu Gly Leu Asp Lys Val Ala Asp Ser
270          165          170          175
272 Lys Val Gly Thr Gln Phe Ile Arg Gly Val Ser Gly Gly Glu Arg Lys
273          180          185          190

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/856,927

DATE: 10/29/2001

TIME: 13:21:11

Input Set : A:\Nih382-1.app

Output Set: N:\CRF3\10292001\I856927.raw

```

275 Arg Thr Ser Ile Gly Met Glu Leu Ile Thr Asp Pro Ser Ile Leu Ser
276      195      200      205
278 Leu Asp Glu Pro Thr Thr Gly Leu Asp Ser Ser Thr Ala Asn Ala Val
279      210      215      220
281 Leu Leu Leu Leu Lys Arg Met Ser Lys Gln Gly Arg Thr Ile Ile Phe
282 225      230      235      240
284 Ser Ile His Gln Pro Arg Tyr Ser Ile Phe Lys Leu Phe Asp Ser Leu
285      245      250      255
287 Thr Leu Leu Ala Ser Gly Arg Leu Met Phe His Gly Pro Ala Gln Glu
288      260      265      270
290 Ala Leu Gly Tyr Phe Glu Ser Ala Gly Tyr His Cys Glu Ala Tyr Asn
291      275      280      285
293 Asn Pro Ala Asp Phe Phe Leu Asp Ile Ile Asn Gly Asp Ser Thr Ala
294      290      295      300
296 Val Ala Leu Asn Arg Glu Asp Phe Lys Ala Thr Glu Ile Ile Glu
297 305      310      315      320
299 Pro Ser Lys Gln Asp Lys Pro Leu Ile Glu Lys Leu Ala Glu Ile Tyr
300      325      330      335
302 Val Asn Ser Ser Phe Tyr Lys Glu Thr Lys Ala Glu Leu His Gln Leu
303      340      345      350
305 Ser Gly Gly Glu Lys Lys Lys Lys Ile Thr Val Phe Lys Glu Ile Ser
306      355      360      365
308 Tyr Thr Thr Ser Phe Cys His Gln Leu Arg Trp Val Ser Lys Arg Ser
309      370      375      380
311 Phe Lys Asn Leu Leu Gly Asn Pro Gln Ala Ser Ile Ala Gln Ile Ile
312 385      390      395      400
314 Val Thr Val Val Leu Gly Leu Val Ile Gly Ala Ile Tyr Phe Gly Leu
315      405      410      415
317 Lys Asn Asp Ser Thr Gly Ile Gln Asn Arg Ala Gly Val Leu Phe Phe
318      420      425      430
320 Leu Thr Thr Asn Gln Cys Phe Ser Val Ser Ala Val Glu Leu Phe
321      435      440      445
323 Val Val Glu Lys Lys Leu Phe Ile His Glu Tyr Ile Ser Gly Tyr Tyr
324      450      455      460
326 Arg Val Ser Ser Tyr Phe Leu Gly Lys Leu Leu Ser Asp Leu Leu Pro
327 465      470      475      480
329 Met Arg Met Leu Pro Ser Ile Ile Phe Thr Cys Ile Val Tyr Phe Met
330      485      490      495
332 Leu Gly Leu Lys Pro Lys Ala Asp Ala Phe Phe Val Met Met Phe Thr
333      500      505      510
335 Leu Met Met Val Ala Tyr Ser Ala Ser Ser Met Ala Leu Ala Ile Ala
336      515      520      525
338 Ala Gly Gln Ser Val Val Ser Val Ala Thr Leu Leu Met Thr Ile Cys
339      530      535      540
341 Phe Val Phe Met Met Ile Phe Ser Gly Leu Leu Val Asn Leu Thr Thr
342 545      550      555      560
344 Ile Ala Ser Trp Leu Ser Trp Leu Gln Tyr Phe Ser Ile Pro Arg Tyr
345      565      570      575
347 Gly Phe Thr Ala Leu Gln His Asn Glu Phe Leu Gly Gln Asn Phe Cys

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/856,927

DATE: 10/29/2001

TIME: 13:21:12

Input Set : A:\Nih382-1.app

Output Set: N:\CRF3\10292001\I856927.raw